

Improving ETMS Default Route Assignment

Aron Futer
CSC/ARCON

January 2005

**Volpe National Transportation Systems Center
55 Broadway
Cambridge, MA 02142**

Executive Summary

Twenty-four hours before a scheduled flight departs, data on this flight from the Official Airline Guide (OAG) is loaded into the Enhanced Traffic Management System (ETMS). This flight is then included in the Monitor/Alert demand predictions that ETMS makes and presents to traffic managers. Since the OAG does not provide the route of flight, ETMS's trajectory model uses a default route, which currently it generates by using the route that this flight has recently flown most frequently. These ETMS historical, default routes, which are used until the flight plan has been received, have been criticized as inaccurate, especially for flights longer than two hours. Since this inaccuracy degrades the predictions of sector loading, this is seen as a serious problem.

The objective of this study is to develop alternatives to the current method and to quantify the improvements that these alternatives could make in the accuracy of predicting routes.

The first alternative uses the route that minimizes the time en route, given the day's predicted winds. This alternative was found to perform worse than the current method.

The second alternative involves recomputing the historical routes every day instead of once a week as currently done. This is called the sliding FS method. It was found that this alternative provided a significant improvement. For example, for one sample of 112,684 flights studied of more than 500 miles in length, the percentage of flights for which the route was correctly predicted rose from the current value of 68 percent to 75 percent.

The third alternative predicts the route for a later flight by using the route in a flight plan for an earlier flight, where the earlier flight must match the later in terms of airline, city pair, and aircraft type. This is called the previous FZ method. By using a previous FZ if it is available and using the sliding FS otherwise, the percentage of flights of more than 2 hours that are correctly predicted rises to 77 percent if one looks three hours before departure or to 76 percent if one looks six hours before departure. Moreover, this method gives especially large increases for the longest haul flights, which are hardest to predict. For example, for flights greater than 5 hours, the percentage of routes correctly predicted is 40 percent for the current method, 45 percent for the sliding FS method, and 52 percent three hours before departure for the previous FZ method.

This report makes two main recommendations.

1. Implement the sliding FS method as soon as possible since this is a relatively small change to ETMS and promises a large increase in accuracy.
2. Implement the previous FZ method on top of the sliding FS method since this promises added accuracy, especially for the flights that are most problematic.

TABLE OF CONTENTS

1. Background..... 4

2. Objective 5

3. Default Route Selection Problem 6

3.1. Airline Route Selection 6

3.2. Current ETMS Approach: Most Popular Historical Routes Used as Default Routes..... 6

3.3. Route Predictions for Longer Flights..... 7

3.4. Searching for Better Approach to Model User’s Route Selection 8

3.5. Collecting Data for Analysis..... 9

4. Results and Discussion 10

4.1. Current Approach vs. Least ETE..... 10

4.2. Weekly Updates vs. Daily Updates 12

4.3. Previous FZ Method 13

5. Recommendations 16

6. References. 16

7. Acknowledgements..... 17

Appendix A: Combined Approach to Predicting Routes..... 18

Appendix B: Time-in-Sector Metric..... 19

Appendix C: Algorithm to Calculate Distance between Routes..... 20

Appendix D: Data by Airlines and Regions..... 21

Appendix E: Alternative Ways to Select Default Routes..... 24

Appendix F: Ideas and Thoughts Not Implemented in the Analysis but of Possible Use in a Future Effort..... 25

1. Background

The Monitor/Alert function of ETMS predicts traffic demands for twenty-four hours into the future for airports, sectors, and fixes. A critical piece of data used by ETMS to make these predictions is the route that the flight will fly. If a flight plan has been filed, then ETMS uses the route in the flight plan. If a flight plan has not been filed, then ETMS must use a default route that is determined by ETMS. The FAA and airlines have expressed unhappiness with the low accuracy of these default routes, especially for long haul flights. Therefore, the purpose of this study is to investigate possible improvements to the way that ETMS determines its default routes. This report is only concerned with flights in the Official Airline Guide (OAG), and therefore only concerned with flights of air carriers, since these are usually the only flights for which a default route is needed.

A scheduled flight first enters the ETMS database 24 hours before the scheduled time of departure. (Prior to the deployment of ETMS 7.9 in November of 2004, a scheduled flight entered ETMS 15 hours before departure.) The OAG data available at this time contains flight ID, origin, destination, and aircraft type, but it does not contain flight path, cruising speed, or cruising altitudes. Currently, ETMS assigns data for these three fields (in a way to be described below) based on historical data for the same carrier, city pair, and aircraft type. The combined OAG and historical data for a flight enter ETMS in what is called an FS message, and this flight is then included in Monitor/Alert's traffic demand predictions. The flight path is commonly referred to as the "field 10" since it is the tenth field in a flight plan.

ETMS uses the historical route for making Monitor/Alert predictions until the flight plan (FZ message) is received. FZ messages are similar to FS messages, but an FZ contains the NAS user's intended flight path, cruising speed, and cruising altitude; since the data in the flight plan represents what the NAS user intends to fly that day for that particular flight, this data is more accurate than the ETMS historical data. Typically, FZs enter ETMS one to two hours before scheduled departure times. ETMS replaces the historical flight data received from the FS messages with the filed flight data from the FZ messages.

Several analyses have been conducted to study the problem of accuracy of ETMS's default route selection. A 1992 analysis carried out at Volpe preceded the initial development of the function that used historical routes with FS messages filed with ETMS. It showed a 94 percent match between fields 10 of the FS and FZ messages when flights were grouped by city pair and airline; city pair alone was successful for 65 percent of the flights (see Reference 1).

A Volpe study by Kip Brown (1999, see Reference 2) found overall match rates at field 10 level at about 80 percent. A Metron Aviation study (Mark Klopfenstein et al, 2001, see Reference 3) used a different (and more liberal) metric to compare routes: the FS route is defined as similar to the FZ route if lists of sectors traversed by these routes match each other. The study found a match in 72 percent of all flights, while for flights longer than 2 hours a match was found only in about 38 percent of the flights.

We must be cautious when comparing these different studies because of differences in methodology and flight samples. At the same time the deterioration of default route predictive

quality over the years may be a sign of increasing sophistication of airline operations, especially in flight planning.

Both traffic management specialists and airline representatives have pointed out that ETMS's historical routes often are wrong, especially for long haul flights. This degrades the Monitor/Alert predictions, especially of sector loading. This appears to be a serious problem, so improving the accuracy of the default routes would improve the usefulness of Monitor/Alert. Introducing Reroute Advisory routes into ETMS (currently under development) may help to improve route predictability but might not be enough.

Monitor/Alert data is used to make demand projections for airports, sectors, and flow evaluation areas/flow constrained areas (FEA/FCAs). The most important time frame for demand projection is usually the next 6 to 8 hours. NAS users submit flight plan messages one to two hours ahead of flight departure, thus a time gap exists when ETMS does not have reliable flight route data from the NAS users. Early intent messages submitted by airlines would alleviate this problem. These messages are not yet widely available, however, so an approach is needed that does not depend on having early intent.

The ultimate goal of the analysis effort is to improve Monitor/Alert performance. However, default route prediction quality is responsible for only a certain part of Monitor/Alert data. This is especially true when one remembers that the suggested improvements target mostly longer flights. The flights longer than 2 hours constitute 20-25 percent of the overall number of flights and are responsible for (roughly) 40-45 percent of the overall en-route sector occupancy. Other possible improvements may include the use of adaptation files for better trajectories on the ascent and descent phases, better models with regard to altitude and speed profiling, and others that are all beyond the scope of this analysis.

2. Objective

The goals of this analysis are as follows.

- Understand the accuracy of the current default routes.
- Develop alternatives to the current default routes.
- Use ETMS data to quantitatively compare the current default routes with the alternatives and determine the potential improvements that could be made.
- Make recommendations on what ETMS should use for default routes and on other activities in this area.

Various methods of selecting default routes for ETMS are presented and evaluated. The best of them are recommended for implementation.

3. Default Route Selection Problem

3.1. Airline Route Selection

Prior to receiving route data from an airline for a flight, ETMS needs to predict the route that the airline will file for that flight. Among the factors that influence airline route selection are the following (quoted from Reference 4).

- Arriving on time (from the airline's perspective), including:
 - Passenger connections
 - Destination gate and manpower planning
 - Next assignment for airplane and crew
- Airplane operating costs
- Fuel costs
- Aircraft weight
- Weather (winds, thunderstorms, turbulence, mountain waves...)
- Desire to avoid overcrowded routes
- Priority changes due to irregular operations

It must be stressed, however, that ETMS does not have specific knowledge of each airline's business priorities, and ETMS does not have specific data on each flight such as how heavily the aircraft is loaded. ETMS can only predict default routes based on the data that it has. At the beginning of this study, one alternative considered was to acquire airline flight planning software and to use it to predict routes. Discussions with airline personnel, however, led to the conclusion that ETMS does not have access either to the values of the parameters used by each airline or to the flight-specific information that is fed into flight planning software; it was judged that this was not a fruitful approach.

This led to us to ask what is the best approximation that ETMS can make using the data that is available, and this led to the least ETE approach described below, which was the beginning point of this study.

3.2. Current ETMS Approach: Most Popular Historical Routes Used as Default Routes

ETMS currently chooses its default route by looking at the most popular route filed over the previous week for each city pair and airline and assuming that this route will be flown. These routes are often called historical routes. (Historical routes, default routes, and FS routes currently are synonyms.)

In more detail, default routes are currently chosen in the following way. When an air carrier or air taxi flight plan is received, ETMS saves the following data:

- flight ID
- aircraft type
- origin and destination

- route of flight
- speeds and altitudes

Once a week (early Wednesday morning), ETMS processes the flight plan data collected over the previous week and for each airline and city pair selects the most commonly used route as the historical route. Sometimes two routes are chosen that differ by aircraft type. Then, when a specific flight is fed from the OAG into ETMS twenty-four hours before departure, the relevant historical route is used for trajectory modeling. See Reference 1 for more details on current historical route processing.

This approach seems to work sufficiently well for short (shorter than 1 hour) and short-medium (between 1 and 2 hours) flights. However, the longer the route, the more choices airlines have, and the harder it is for ETMS to predict what the choice would be.

3.3. Route Predictions for Longer Flights

A study by Metron (Reference 3) provides the following statistics. The study presents data collected on 18 days in June 2000 for roughly 446,000 flights. It compares ETMS historical routes (FS) to airline filed routes (FZ). Routes are considered to be similar if they are projected to pass through the same set of sectors. This study's main results are as follows.

- For roughly 170,000 flights shorter than 1 hour: 90.5 percent of them had similar FS and FZ routes;
- For roughly 176,000 flights between 1 and 2 hours long: 72.5 percent of them had similar FS and FZ routes;
- For roughly 99,000 flights longer than 2 hours: 38.5 percent of them had similar FS and FZ routes.

These results indicate a fundamental finding that appears to be generally true, which is that the longer the flight, the harder it is to predict the route. The explanation for this appears to be that because there are so many more route choices for longer flights, it is harder to predict the specific route that will be filed for a particular flight.

To see why it might be difficult to predict the route that will be filed for a long haul flight, consider the situation shown in Figure 1. The data reflects the route choice by two airlines on a cross-country route for flights filed for a city pair for the same time period (45 days). Airline A had 90 different routes for 116 flights filed during this time; 56 routes were used only once representing about 48 percent of flights; the maximum number of flights using the same route was 12, which represented about 7 percent of the flights. Airline B had 19 different routes for its 160 flights filed during the same time; 10 routes were used only once representing about 6 percent of flights; maximum number of flights using the same route was 98, which represented about 61 percent of flights. This shows that for airline B, it is fairly easy to do a good job of predicting the route, but it is very difficult for airline A.

The routes were compared by field 10 texts; however the routes were widespread geographically, too, when shown on the US map.

Our discussions with airline personnel indicate that in the future the fragmented route choice exhibited by airline A might well become the more typical mode of operation.

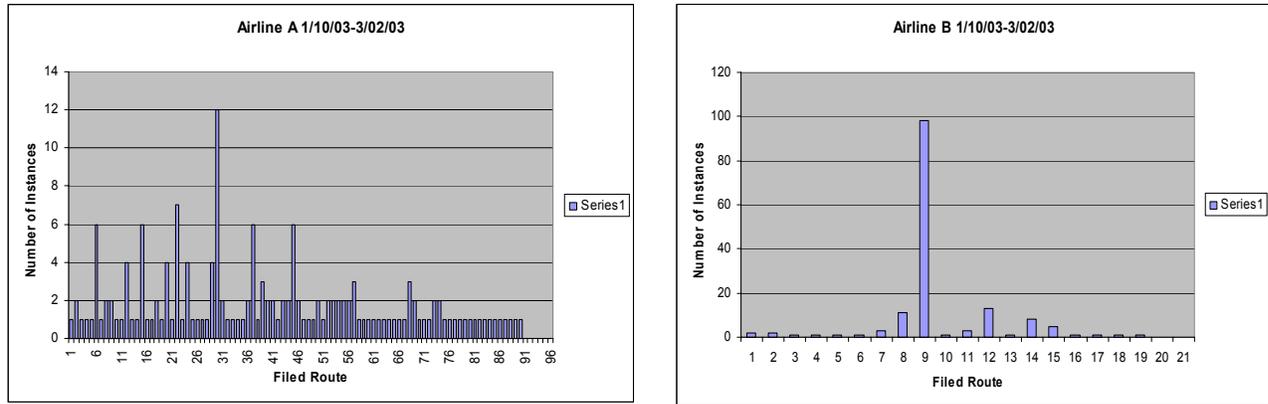


Figure 1. Based on Airline Operations, How Much Can Be Done to Improve Route Predictions?

3.4. Searching for a Better Approach to Model User’s Route Selection

For the current analysis effort, several methods to select default routes were proposed, tested, and analyzed.

For one airline, one city pair, and a specific aircraft type, let us consider a route candidate pool as composed of routes filed in the past by the airline, the city pair, and the aircraft type. Then let us consider the following methods of route selection.

- Frequency of use (*FS* method): Of the route candidates, use the most popular one filed during the previous week; this is the current method used to select a default route.
- Least estimated time en route (*LETE* method): Of the route candidates, use the ETMS trajectory model to pick the least-time route, given the day’s winds.
- Recency of previous use (*Previous FZ* method): Use flight plans (*FZ* messages) filed earlier in the day by an airline to predict routes filed by the same airline later in the day.

These methods will be described and evaluated below in Section 4. An additional method that combines these three methods is covered in Appendix A. If fully implemented it might have potential benefits.

See Appendix E for a discussion of alternate ways to select default routes. See Appendix F for a collection of ideas that were not pursued in this study but that might be taken into consideration in a future study.

3.5. Collecting Data for Analysis

For the analysis, only commercial flights were used. (GA flights are not present in OAG and thus FS messages are not filed for them.) The set of flights was limited to those between 80 major US airports. The data that is analyzed in this study was collected for two periods. First, for 15 days in the summer of 2003, data was gathered for about 47,000 flights that were longer than 2 hours. (Shorter flights were omitted since the current method already does a good job for shorted flights.) Second, for 20 days in the winter of 2004 (i.e., the first few months of calendar 2004) data was gathered for about 113,000 flights that were longer than 500 miles.

ETMS submits FS message for scheduled flights 24 hours prior to departure (when this analysis was conducted this value was 15 hours). ETMS receives updated wind forecasts every 3 hours. For the analysis, all the routes filed by each airline for each city pair and aircraft type were stored. For each flight these routes were later “filed” as fake FS messages with every wind update between the original FS message and the departure, 5 times total. This provided the estimated time en route (ETE) from the ETMS trajectory model as well as the list of sectors that the routes traversed. The number of routes filed for each flight varied widely, say, from 10 to 100, with the average around 20.

It was confirmed in the course of the analysis that the predicted winds change during the day but very slowly, and so changes in the wind projections do not seem to appreciably affect route choice over the course of one day.

The summer 2003 effort was done with the use of an ETMS development string and post-processing of the so-called *orig* files. However, “filing” the messages with all these routes every three hours and running them through the ETMS trajectory modeling imposed a substantial burden on the system and was very labor-intensive. This method worked, though a small amount of data was lost. A different approach was used to analyze the data for winter 2004; a stand-alone customized ETMS mini-string was used. This approach proved to be much faster and also handled the data more reliably.

The flight data was grouped by length of flight.

Various metrics were considered to measure the quality of route prediction, namely various ways to estimate the percentage of flights for which the filed route was correctly predicted.

- Literal textual match between field 10 of a default route and field 10 of a FZ route.
- “Distance” between a default route and FZ route: 0, 5, 10, 20, 50, 100 miles; (see Appendix C for the distance algorithm).
- List of sectors traversed:
 - complete sector list match between a default route and FZ route;
 - sector list match for all sectors with exclusion of those around departure and arrival airports;
 - 80 percent of all sectors match;
 - 80 percent match of all sectors with exclusion.

These metrics were applied to various data samples. In most cases, different metrics would produce different percentages of flights for which FS and FZ routes would match. The literal match of field 10s was not an appealing metric since different field 10s can represent exactly the same route. Examination of the performance of these metrics led to the conclusion that the results do not depend sensitively on which metric is used. We decided to use only one metric, complete sector list match between a default route and FZ route, the same metric used by the Metron study (Reference 3).

Late in the study one more metric was introduced: percentage of time spent by all the flights in the proper sectors. This metric better reflects actual sector loads used in Monitor/Alert. It is discussed in the Appendix B.

4. Results and Discussion

4.1. Current Approach vs. Least ETE

The first alternative to the current approach to be considered is the least estimated time en route (LETE) approach. The rationale for this approach is that a good candidate for prediction for the route that will be filed is the route that, according to the calculations of the ETMS trajectory model, yields the least time in the air of all the recently filed routes. The LETE route can be considered to be an approximation to the least cost route. In fact, the original rationale for this study was that airline personnel pointed out that the ETMS historical routes sometimes led to ETMS assuming a route that was flying right into the jet stream and that was a very poor estimate of the route for that day.

Table 1 compares the baseline (current approach, FS method) with the least ETE method. In this table, as well as in other tables below, the data for summer 2003 and winter 2004 is presented separately. Each row in this table presents data on a set of flights that are grouped by flight length. For concreteness, this table will be explained by explaining one particular row.

Consider the row under winter 2004 for flights that were between 3 and 4 hours in length. It is seen that there were 17,806 of these flights, which represent 24 percent of the 74,906 flights that lasted more than two hours. The FS method, which is the current method, correctly predicted the route for 58 percent of these flights. The LETE method correctly predicted the route in 42 percent of these flights. Other rows are interpreted analogously. In this table and others the percentage correctly predicted is calculated using the sector-list metric; that is, two routes are considered the same if the list of sectors traversed by the two routes are exactly the same.

The following conclusions can be drawn from the table.

Season	Flight Length	# Flights	Percentage Correctly Predicted	
			FS	LETE
Winter 2004	> 500 miles	112,584	76	61
	< 2 hours	37,678	90	77
	> 2 hours	74,906	68	53
	2-3 hours	44,585	79	63
	3-4 hours	17,806	58	42
	4-5 hours	8,069	47	30
	> 5 hours	4,446	40	28
Summer 2003	> 2 hours	47,348	53	38
	2-3 hours	27,666	62	48
	3-4 hours	12,062	45	29
	4-5 hours	5,673	32	18
	> 5 hours	1,947	25	17

Table 1. Current Approach vs. LETE.

First, the results vary significantly by season. For example, for flights between 3 and 4 hours, the FS method provided the correct route in the winter 2004 in 58 percent of the cases, while in summer 2003 in 45 percent of the cases. We conjecture that the explanation of better predictions in the wintertime is that the weather is more stable in the winter.

Second, the current method provides very good results for short flights. This table only covers flights that are more than 500 miles; in winter 2004 there were 37,678 flights of more than 500 miles but with an ETE of less than 2 hours, and the FS method correctly predicted the route for 90 percent of these flights. This means that it is for longer flights that better default routes are needed. (Note: It is worth mentioning that the number of really long flights (the last two categories in each group, 4-5 hours long and longer than 5 hours) is small comparing to overall number of flights. This means that even with lower predictability of default routes for these flight groups their impact on the overall Monitor/Alert should be limited.)

Third, the FS method is consistently better than LETE. This finding at first was a surprise because the FS method does not take into account the winds on the day of flights and the LETE method does. We had conjectured that using the knowledge of the day's winds should have been a clear winner. After a careful analysis, however, we determined the apparent reason for this initially surprising result.

The reason is that in the distribution of the ETEs for the route candidates that had been run through the ETMS trajectory model the lower end is always crowded. This means that out of, say, 20 route candidates, it were often two or more routes with the same least ETE, plus possibly several more with ETE only a few minutes greater. So several "best" routes competed for the first spot, and the LETE method could not always choose the same one that the airline would later file. Also, several minutes' difference in ETE estimation could have easily been explained by difference in wind model of the airlines and ETMS. The conclusion is that using the least

ETE method alone is a poorer predictor than using the FSs, but, as discussed in Appendix A, there might still some use that could be made of LETE in combination with other methods.

Note: The baseline figure for FS route match for flights longer than 2 hours in summer 2003 is 53 percent. Comparable figure in Metron study taken in June of 2000 (Reference 3) is 38.5 percent. We do not have an explanation for this discrepancy other than changes in traffic and weather patterns.

4.2. Weekly Updates vs. Daily Updates

The database of historical routes is currently updated once a week, namely every Wednesday. To see how this can result in stale data being used, consider an example. Early in the morning of Wednesday, December 15, new historical routes were calculated using data for the week December 8-14. This means that when these new historical routes began to be used on December 15, these routes were based on data from December 8-14 that was from 1 to 7 days old; these routes were still being used on Tuesday, December 21, at which time these routes were still based on data from December 8-14 that was now from 7 to 13 days old. In other words, historical routes were being used on December 21 that were based on data that was needlessly old. This study looked into the question of whether it mattered that needlessly old data was being used.

In the current system, new historical routes are created once a week. As an alternative, assume that new historical routes are still based on a week's worth of data but that they are created every day. For example, under this alternative, on December 15, historical routes being used would have been based on data from December 8-14, just as in the current system; on December 21, in contrast, the historical routes would have been based on data from December 14-20. This method of calculating new historical routes every day is called the sliding FS method. The rationale for the sliding FS method is that 6 days a week the default routes that are chosen will be based on more recent data, and this leads to the conjecture that they might be better.

Table 2 compares the baseline (current approach, FS method with weekly updates) with the sliding method in which new historical routes are created on a daily basis. The data shows that the sliding method is better across the board, for all seasons and all flight groups.

In particular, for flights in the winter 2004 of between 3 and 4 hours, the weekly build of historical routes gives accurate predictions 58 percent of the time, whereas the daily build gives accurate predictions 65 percent of the time. This is an absolute improvement of 7 percent. If we express this improvement in relative terms, 65 percent is 12 percent greater than 58. If one looks at either the absolute or the relative columns of Table 2, one sees that the sliding method is better in every case, and the improvements are largest for the longer flights, which are the hardest to predict.

As can be seen from the table, the absolute improvement demonstrated by the sliding method is mostly about 5-7 percent. This means that the percentage of flights where FZ routes are correctly predicted by FS routes is about 5-7 percent higher for the sliding method.

			Percentage Correctly Predicted		Percent of Improvement	
Season	Flight Length	# Flights	FS Weekly (batch)	FS Daily (sliding)	Absolute (sliding – batch)	Relative ((sliding – batch) / batch)
Winter 2004	> 500 miles	112,584	76	81	5	7
	< 2 hours	37,678	90	94	4	4
	> 2 hours	74,906	68	75	7	10
	2-3 hours	44,585	79	86	7	9
	3-4 hours	17,806	58	65	7	12
	4-5 hours	8,069	47	52	5	11
	> 5 hours	4,446	40	45	5	12
Summer 2003	> 2 hours	47,348	53	58	5	9
	2-3 hours	27,666	62	68	6	10
	3-4 hours	12,062	45	49	4	9
	4-5 hours	5,673	32	35	3	9
	> 5 hours	1,947	25	27	2	8

Table 2. Weekly Updates vs. Daily Updates

It is smaller – 4 percent – for winter 2004 flights that are shorter than 2 hours because the baseline is rather high – 90 percent, there is not much room for improvement. It is also small – 2-3 percent – for the summer 2003 flights that are longer than 4 hours for the very opposite reason: the baseline data values are small – 25 percent and 32 percent.

However, this perception changes if to look at relative improvement. This is why the last column of the table is included. As is seen from the table, relative improvement by the sliding method for most of the flight groups is between 7 percent and 12 percent. It is 8 percent or higher for flights longer than 2 hours.

We tried variations of the sliding window size: two weeks, four weeks, three days. It turned out that one week showed the best results.

4.3. Previous FZ Method

Consider a flight between a city pair flown by a particular airline with a particular aircraft type. If that airline has flown another flight between the same city pair with the same aircraft type a few hours before the flight in question, then the route filed for the first flight can be used as a prediction of the route that will be filed by the second flight. This is called the previous FZ method since the route from the FZ for a previous flight is used to predict a route for another flight.

The route in a previous FZ can be thought of as a sort of early intent. Since we have found that winds change slowly over the course of a day, the route that an airline actually filed for one flight can be taken as a rough statement of its intent for a succeeding flight.

The previous FZ method can be summarized in the following algorithm that describes the default route that ETMS will use for a particular flight.

1. If an FZ has been filed for an earlier flight between the same city pair, flown by the same airline, and with the same aircraft type within the 24 hours preceding the departure time, then use the route in the most recently filed FZ as the ETMS default route.
2. Otherwise, use the FS route (and preferably use the sliding FS route).

In other words, the algorithm is to use the previous FZ if there is one; if there isn't, then use the FS route. Table 3 shows the basic results for the previous FZ method. Data for this analysis is only available for winter 2004.

To see how this table is interpreted, consider the flights between 3 and 4 hours in length. Fifteen hours before departure, there will be some of these flights for which there is no previous FZ within 24 hours of departure, so the FS route is used; for other flights, however, there will be one or more previous FZs, and so in this case the route from the most recent previous FZ is used. This method results 15 hours before departure in the correct prediction of the filed route in 63 percent of the cases. If we then jump ahead to 12 hours before departure, there will be some previous FZs that come in during the three hours between 12 and 15 hours before departure; these are used to predict the route, and this is enough to raise the percentage of correct routes to 64 percent. Similarly, over time more FZs are received, and these gradually raise the percentage of correct routes. Three hours before departure, ETMS will have correct routes for 69 percent of the flights.

Flight Length	Percentage Correctly Predicted				
	15 Hours	12 Hours	9 Hours	6 Hours	3 Hours
> 500 miles	81	81	81	82	83
< 2 hours	95	95	95	95	95
> 2 hours	74	74	74	76	77
2-3 hours	85	85	85	86	87
3-4 hours	63	64	64	66	69
4-5 hours	51	51	52	54	58
> 5 hours	43	44	46	48	52

Table 3. Previous FZ Results Depending on Lead Time to Departure (Winter 2004).

Two conclusions can be drawn from Table 3. First, the longer the flights, the poorer the predictions are; the same was true in the case of Table 1 and Table 2.

Second, the closer the previous flight is (in time of departure) to the actual flight the better the performance of the previous FZ method.

To make these results stand out more clearly, consider Table 4, which is the same as Table 3 but with two additions. First, the last column shows the percentage correctly predicted by the sliding

FS method; these numbers are taken from Table 2. Second, cells are shaded that show the earliest time that the previous FZ method gives results that are as good as or better than the sliding FS method. For example, for flights between 3 and 4 hours in length, the percentage correctly predicted by the sliding FS method is 65 percent. For the previous FZ method, 15 hours before departure the percentage correctly predicted is 63 percent, and this rises to 64 percent at 12 hours before departure and stays at 64 percent at 9 hours before departure. For later times, the previous FZ method provides better results than the sliding method.

One can draw several conclusions from Table 4.

Flight Length	Percentage Correctly Predicted					
	15 Hours	12 Hours	9 Hours	6 Hours	3 Hours	Sliding FS
> 500 miles	81	81	81	82	83	81
< 2 hours	95	95	95	95	95	94
> 2 hours	74	74	74	76	77	75
2-3 hours	85	85	85	86	87	86
3-4 hours	63	64	64	66	69	65
4-5 hours	51	51	52	54	58	52
> 5 hours	43	44	46	48	52	45

Table 4. Previous FZ (with Lead Times to Departure) vs. Daily (Sliding) FS (Winter 2004)

First, for the flights with routes longer than 500 miles and the flying time shorter than 2 hours the previous FZ method is better than the sliding FS method. Though both numbers are high – in the mid-nineties – and the margin of difference is only 1 percent, it is worth mentioning. The reason for this difference might be that these flights in general are scheduled to fly more often and thus a previous flight route works like the early intent route. Also, for flights shorter than 2 hours the route choice is not that wide as for longer flights.

Second, for the flights that are shorter than 500 miles and were not a subject of this study it might as well be beneficial to switch from the FS to Previous FZ method. This suggestion would need to be separately investigated.

Third, for flights longer than 2 hours it is seen that moving through time, the earlier previous FZ predictions are worse than sliding FS predictions, but the later Previous FZ predictions are better.

From the third conclusion the following strategy for ETMS default route predictions may be derived.

- For a flight that is projected to fly less than 2 hours use Previous FZ method (at least when the previous route is available within reasonable time frame).
- If a flight is projected to fly longer than 2 hours, start with sliding FS route when FS message is submitted. And then switch to available Previous FZ route around 6-9

hours prior to the flight's departure. Update the predicted route with each newly available Previous FZ.

This strategy will not improve Monitor/Alert sector loading data for all time intervals, but will improve predictions for the time interval that is most crucial for purposes of traffic flow management.

A simpler strategy with nearly the same benefit might be to start with the FS route in every case but to switch to the previous FZ at 6 hours prior to departure.

Further research may be needed to better investigate and then adjust the times of switching from one method to another. It might be that the best results will be achieved by a flexible self-learning strategy that adapts over time based on the recent performance of various methods. As a part of that, it may be that for some sub-groups of the 2-3 hour group Previous FZ method might be better for all time buckets. Identifying such sub-groups (say, by airlines, time of the day, city pairs, etc.) was beyond the scope of this study. Appendix D gives some insight on the differences in predictability for different categories of flights.

For 4-5 hour and longer flights the border cells (which are colored in Table 4) appear earlier than for 2-3 and 3-4 hour flights. The reason for that might be that the longer routes depend most strongly on the winds, and the projected winds change slowly over time. So these are the flights for which Previous FZ provides the most information. (Interestingly, for the flights shorter than 2 hours Previous FZ is also good predictor, but for a different reason).

5. Recommendations

The analysis leads to the following recommendations:

1. Implement the sliding FS method as soon as possible. It looks like this is reasonably easy to implement, and this is the most beneficial.
2. Implement the method that combines Previous FZ with sliding FS. This definitely brings improvements, but this will be considerably more complicated to implement than the sliding FS method.
3. Create software to permanently monitor Monitor/Alert performance, quality of route predictions.
4. Continue working with airlines to convince them to file flight routes earlier. (The airlines often complete flight planning 2-3 hours prior to the flight; for long flights this may be good enough to cover substantial portion of the flight. Nightshift dispatchers may do pre-planning.)
5. Revisit analysis to further improve route predictions in the future.

6. References.

1. "Determination of Flight Plan Data for Scheduled Flights", ETMS Functional Description Version 7.6, Appendix A, August 2003.

2. “Assessment of the Routes Assigned to the FS Messages”, Kip Brown (Volpe Center), report on analysis conducted in support of ETMS Monitor/Alert task, June 1999.
3. “En Route Automation Modernization; Flight Plan Preprocessing Prototype Project: Analysis of Early Intent Flight Planning and the Predictability of Sector Demand”, Final Report by Bruce E. Ware, Mark Klopfenstein, and Andrew Ryan, Metron Aviation, July 19, 2002.
4. “The Role of Free Flight in Delta’s Future” by G.F. Bell and R.E Curry (Delta) at the AIAA Transportation TC Workshop, Washington, DC, March 06, 1996.

7. Acknowledgements

Barry Davis (FAA) initiated the project. During the course of the project, Rick Oiesen and Ken Howard always helped and gave advice. Bob Hoffman, Mike Brennan, Mark Klopfenstein (all from Metron) helped a lot with discussion and advice. Two presentations made at the Metron seminars in December 2003 and July 2004 lead to valuable discussions with priceless feedback. Bing Zheng programmed the bulk of the analysis software, with the help of Bob Sharick, Sylvia Toderò, and Boris Epstein.

Appendix A: Combined Approach to Predicting Routes

The algorithm described in this appendix attempts to combine the best elements of the various approaches described in Section 4. This appendix describes how one could select routes that were flown recently, appeared in the route candidate pool at least several times, and would be not much longer than the least-ETE routes. Several ways to implement this idea were possible, the following one was chosen.

For an airline, city pair, and aircraft type, the routes that were filed during the previous week were considered, and among them only those that were filed more than once were selected. Then out of the selected routes the ones that were filed through the same set of sectors were grouped together. Then the routes with ETEs that were significantly longer than the least ETE for the group were discarded. Finally, if the remaining set contained the previously filed route it was chosen; otherwise the most frequently filed one was chosen.

In other words, this approach first throws out all routes that, given today's winds, are inappropriate. Then, from the set of appropriate routes it chooses the route that appears in the previous FZ if available and otherwise the route that has recently been flown most often.

Table 5 shows the results. The results in this table were calculated in a somewhat different way from the tables above. The route predicted by the sliding FS method is not time dependent, but the route predicted by the previous FZ method does depend on how far ahead departure one looks because new flight plans can be received as time passes; since the combined approach depends in part on the previous FZ approach, the combined approach is also time dependent. The time used to calculate the results in Table 5 was the time of the wind update that occurred between 3 and 6 hours before departure.

	Flight Length	Sliding FS	Previous FZ	Combined
Winter 2004	> 500 miles	81	83	83
	< 2 hours	94	96	94
	> 2 hours	75	78	77
	2-3 hours	86	87	87
	3-4 hours	65	69	68
	4-5 hours	52	59	56
	> 5 hours	45	54	48

Table 5. Combined Approach.

The main conclusion is that the combined approach is better than sliding FS approach, but slightly worse than the previous FZ approach. That is, this particular algorithm does not provide an advantage over the previous FZ approach.

Nevertheless, this approach is worth considering because of its major benefit: it would not choose a route that would cause a flight to fly across the country right into the jet stream. The disadvantage of this approach is that it is more complicated to implement; to run scores of routes

through the ETMS trajectory modeling for each flight will create additional burden on the system. However, the complexity of this approach is comparable with the complexity of the initial pure LETE idea.

Possibly, more elaborate implementation and better tuning could improve the model performance for the combined approach. However, the question remains though to what extent improvement might be achieved due to the law of diminishing returns.

Appendix B: Time-in-Sector Metric

A question not answered by this analysis is: What is the best absolute way to measure the degree of the correctness of the default route? As it is mentioned above, a number of metrics were considered and finally the sector-list was used during the course of the analysis. This metric proved to be sufficient for a comparative analysis of the default route algorithms. At the same time, however, one more metric deserves to be mentioned; Mike Brennan (Metron) should be credited with the idea.

The new metric, referred to as time-in-sector metric, can be best explained in contrast to the sector list metric. The sector list metric requires an exact match between the sector list for the default route and the sector list for the filed route. Using this metric, a default route that matches 95% of the right sectors is counted the same as a default route that matches none of the right sectors; that is, as a failure. In fact, the 95% matching route contributes much more positively to Monitor/Alert and should receive a much higher grade.

The time-in-sector metric provides a fairer grade by considering the percentage of time that each flight spends in the right sectors, and accumulating these percentages to compute the overall score. Furthermore, the time-in-sector not only considers which sectors match but how much time each flight spends in the right sectors and wrong sectors. When aggregated, the time-in sector metric provides a true indicator of what percentage of time the default routes put flights in the right sectors.

Winter 2004	Time-in-Sector				Sector-List			
	Sliding		Previous	Combined	Sliding		Previous	Combined
	FS	LETE	FZ		FS	LETE	FZ	
> 500 miles	85	81	87	86	81	72	83	83
< 2 hours	98	94	98	98	94	87	96	94
> 2 hours	82	77	84	83	75	65	78	77
2 - 3 hours	92	87	92	92	86	76	87	87
3 - 4 hours	79	73	81	80	65	54	69	68
4 - 5 hours	70	66	75	72	52	43	59	56
> 5 hours	63	61	71	65	45	39	54	48

Table 6. Time-in-Sector vs. Sector-List Metric.

Table 6 presents a comparison of the time-in-sector metric and the sector list metric for different default route algorithms. As one can see, the relative comparisons between the algorithms are all the same: the longer the flights, the have poorer the predictions; the FS method is better than LETE (sliding LETE means LETE-data collected with sliding window as well as for FS-data); the Previous FZ method is better than the FS; finally, the combined method is a little worse than Previous FZ. All the conclusions of the main analysis are unchanged.

However, there are two interesting observations about this metric.

First, what is interesting about this metric is that the drop off in accuracy for longer flights is less than indicated by the sector-list metric. For example, if you compare the FS method across the ETE ranges, the sector-list metric drops from 94% for less than 2 hours to only 45% for greater than 5 hours. Using the time-in-sector metric, the drop off is from 98% to 63%. The most likely reason is that: for flights crossing half of the US traversing long distances, it is much harder to have the entire lists of sectors equal, but many of them might well match. Another observation is that for the time-in-sector metric relative drop off for the Previous FZ method is noticeably smaller than for other methods; no explanation is given for that.

The other observation is that when statements are made about the absolute accuracy of the ETMS default routes, the picture looks much better when using the time-in-sector metric. If we adopt the sliding FS method, for example, and can state that routes are correct 98% of the time for less than 2 hours and 82% of the time for greater than two hours, the accuracy of the routes would appear to be pretty good.

Appendix C: Algorithm to Calculate Distance between Routes

Suppose that we have two routes **R1** and **R2** between two cities. Suppose that each of the routes is defined by sequence of waypoints: (V_1, V_2, \dots) for route **R1**, and (U_1, U_2, \dots) for route **R2**. Each consecutive pair of waypoints (V_i, V_{i+1}) or (U_i, U_{i+1}) defines a segment.

The core of the route distance algorithm is a method that for each point **A** and segment (B, C) on the sphere will calculate the distance $D_s(A, (B, C))$ between **A** and (B, C) . The method is described later in this section.

For each waypoint V_i of route **R1** we will define the distance $D_w(V_i, R_2)$ between V_i and **R2** as minimum by **j** of $D_s(V_i, (U_j, U_{j+1}))$. Then we will define the so-called pre-distance $PD(R_1, R_2)$ as maximum by **k** of $D_w(V_k, R_2)$. And finally, we will define distance $Dr(R_1, R_2)$ between the routes **R1** and **R2** as larger of two values: $PD(R_1, R_2)$ and $PD(R_2, R_1)$.

The algorithm for calculating distance $D_s(A, (B, C))$ between **A** and (B, C) follows.

If angle **ABC** is obtuse $D_s(A, (B, C))$ is defined as the length of the segment of the geodesic line on the Earth's sphere between points **A** and **B**; see Figure 2, left.

Otherwise, if angle \mathbf{BAC} is obtuse $\mathbf{Ds(A, (B, C))}$ is defined as the length of the segment of the geodesic line on the Earth’s sphere between points \mathbf{A} and \mathbf{C} ; see Figure 2, middle.

Otherwise, $\mathbf{Ds(A, (B, C))}$ is defined as the length of perpendicular segment (on the sphere) from point \mathbf{A} to segment $\mathbf{(B, C)}$; see Figure 2, right. This is how it is calculated.

Let us consider two vectors \mathbf{OB} and \mathbf{OC} from the point \mathbf{O} (center of the Earth) to the points \mathbf{B} and \mathbf{C} , respectively. Let us consider vector \mathbf{OG} that is defined as $[\mathbf{OB}, \mathbf{OC}]$ – cross product of vectors \mathbf{OB} and \mathbf{OC} . Now let us define vector \mathbf{ON} as normalized vector \mathbf{OG} – vector \mathbf{OG} divided by its length (vectors \mathbf{OG} and \mathbf{ON} are not shown on the picture). Then the distance between \mathbf{A} and $\mathbf{(B, C)}$ will be equal to $|\langle \mathbf{OA}, \mathbf{ON} \rangle|$ – absolute value of dot product of vectors \mathbf{OA} and \mathbf{ON} .

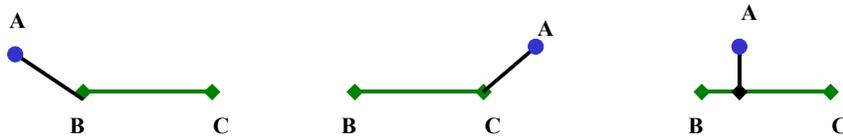


Figure 2. Defining Distance between a Point and a Segment.

Appendix D: Data by Airlines and Regions

Airlines differ in predictability of the routes they file. Table 7 shows the data collected in summer 2003; this covers all of the flights, i.e., the same 47,348 flights reported on in Table 1. The data is shown for eight major airlines separately (American, United, US Airways, Continental, Delta, North West Airlines, South West Airlines, and America West); all the smaller airlines are lumped together as the “Rest”.

	% of traffic	Percentage Correctly Predicted	
		Sliding FS	Previous FZ
All Airlines	100	53	60
Airline 1	24	52	63
Airline 2	15	34	42
Airline 3	12	53	59
Airline 4	10	39	50
Airline 5	9	51	61
Airline 6	6	52	53
Airline 7	4	57	53
Airline 8	5	89	90
Rest	16	72	72

Table 7. Summer 2003 Data by Airlines.

Each row is dedicated to one airline and has cells with: percentage of the nationwide traffic that a specific airline carried, percentage of correctly predicted routes by sliding FS method, and percentage of correctly predicted routes by Previous FZ method. It is visible that the two categories are substantially better in predictability of their routes: Airline 8 and the Rest. The differences are likely attributable to different operational practices, city-pairs served, regions covered, etc. There is no evidence to suggest that this approach would work better for different airlines.

Table 8 shows the data in respect to the region of the country where the origin and destination airports belong. The US was divided into Eastern and Western regions by Mississippi river. EE means that both origin and destination are on the East, WW – on the West, and EW and WE fly from one region to another. As it can be seen, WW traffic is the most predictable, and then follows EE, and the EW & WE traffic is the least predictable. Possible explanations for these phenomena:

- EE and WW traffic is more predictable because in general the North-South and South-North routes are shorter than East-West or West-East routes;
- EE traffic is most often subject to traffic control; this is why WW traffic shows better predictability.

	% of traffic	Percentage Correctly Predicted	
		Sliding FS	Previous FZ
All Regions	100	53	60
EE	22	60	67
WW	16	74	81
EW & WE	62	45	52

Table 8. Summer 2003 Data by Regions.

Another observation is derived if to combine the information by the airlines and by the regions. Table 9 shows the data for Airline 8 and the Rest divided by regions. Now let us combine “good” airlines with “good” regions; namely, consider all of EE and WW traffic in combination with EW and WE traffic of Airline 8 and the Rest. This group of flights constitute about 50 percent of the overall traffic, and for this group the default route predictability is noticeably higher than total.

How this type of considerations might help? A couple of ideas:

- If categories of “bad” flights are smaller than total, may be airlines will be willing to give early intent at least for those flights;
- Predictability for groups of flights may help to measure confidence of Monitor/Alert data.

		% of traffic	# of flights	Percentage Correctly Predicted	
				Sliding FS	Previous FZ
All Regions	Airline 8	5	2246	89	90
EE	Airline 8	3	307	92	88
WW	Airline 8	8	636	89	98
EW & WE	Airline 8	4	1303	89	87
All Regions	Rest	16	7585	71	72
EE	Rest	26	2709	74	77
WW	Rest	19	1423	79	83
EW & WE	Rest	12	3453	66	64

Table 9. Summer 2003 Data by Regions for “Good” Airlines.

Tables 10 and 11 present winter 2004 data dissected by airlines and regions. Most of the conclusions are the same as for the summer 2003 data.

	All flights > 500 miles				All flights > 2 hours			
	% of traffic	Percentage Correctly Predicted			% of traffic	Percentage Correctly Predicted		
		Sliding FS	Previous FZ	Combined		Sliding FS	Previous FZ	Combined
All Airlines	100	81	84	83	100	75	78	77
Airline 1	8	76	82	79	11	72	79	75
Airline 2	14	71	74	72	12	58	62	60
Airline 3	9	81	84	83	6	75	79	78
Airline 4	7	78	80	79	11	74	78	75
Airline 5	14	81	84	83	14	73	76	74
Airline 6	6	63	69	67	8	55	62	59
Airline 7	5	71	73	73	6	65	67	68
Airline 8	11	97	97	97	11	97	96	97
Rest	25	92	92	92	22	87	88	88

Table 10. Winter 2004 Data by Airlines.

	% of traffic	Percentage Correctly Predicted		
		Sliding FS	Previous FZ	Combined
All Regions	100	82	84	83
EE	38	93	93	94
WW	21	88	89	89
EW	21	64	70	66
WE	21	72	75	74

Table 11. Winter 2004 Data by Regions.

Appendix E: Alternative Ways to Select Default Routes

This section lists several ideas for default routes along with their benefits and disadvantages. Some of the approaches have been explored in the analysis and discussed in detail in this report; others were not.

- Of the routes used in the past by an airline for a city pair, use the ETMS trajectory model to pick the least-time route, given the day's winds.
 - Pros:
 - It is feasible since it uses available data and modeling.
 - Takes into account different airlines' models (to an extent).
 - Works well for emulating conservative route selection, i.e. search among the existing routes.
 - Cons:
 - Puts heavy weight on the single factor in flight planning, however other factors might be more important for an airline on a particular day.
 - Does not suggest how to choose between two or more candidate least-ETE routes, thus reducing the chance of the correct guess.
 - Implies that airlines' wind model works similarly to the ETMS's one.
 - Does not work well for emulating dynamic route selection; i.e., it cannot introduce a new route into the model.
- Of the routes used in the past by an airline for a city pair, use the ETMS trajectory model to pick the route closest in ETE to airline's historical ETEs, given the day's winds. Most pros and cons are similar to the least-ETE model. What differs:
 - Pros:
 - Might better reflect airlines' thinking in route selection: use ETE as an input to the model and find the least-fuel-burn route for it;
 - Reduces the guessing game to a single parameter which is much more continuous than routes; averaging concepts might apply that don't work as well for geographical routes.
 - Cons:
 - Determining least-fuel-burn route might not be easy to implement within ETMS;
 - Might mostly suggest the same routes as the least-ETE model, so why bother.
- Use flight plans filed early in the day by an airline to predict routes filed by the same airline later in the day.
 - Pros:
 - Since winds are fairly stable over the course of a single day, this could be quite accurate. It is very feasible.
 - Works well for emulating conservative route selection.

- Cons:
 - This works best when the previous flight for the same airline and city pair had been filed within a few hours.
 - Does not work well for emulating dynamic route selection.
- Assume that flights take similar routes on similar days (weather-wise). For example, if the jet stream is in a particular position on a day, flights will take the same route as on other days when the jet stream is in that position.
 - Pros:
 - This could improve accuracy since it seems to capture most of weather-related factors of flight planning.
 - Cons:
 - Judging when days are “similar” is a big problem. This consideration keeps this from being very feasible.
- Acquire flight planning software.
 - Pros:
 - This might greatly improve accuracy.
 - This might be better tunable to airlines’ priorities than just historical routes.
 - More dynamic: might suggest routes that are not currently in the ETMS database.
 - Easier to keep up with innovations in flight planning.
 - Provides 4-D route planning: route and vertical profile with times would lead to better M/A predictions.
 - Cons:
 - This is an unknown area; we do not know what it might take to integrate it into ETMS.
 - Different airlines use different software and business models. Airlines have special considerations for some flights. Incorporating all of this into ETMS does not look feasible. ETMS simply does not have parameters (e.g., take-off weight) to feed the model. Will we be able to emulate it?
 - It is unproven that the idea will increase accuracy.
 - It might be costly in \$\$.

Appendix F: Ideas and Thoughts Not Implemented in the Analysis but of Possible Use in a Future Effort

This section presents some leftover thoughts and ideas either not explored at all or just partially explored and implemented during the course of the analysis. A lot of this was derived by compilation of quotes from e-mails composed by Rick Oiesen, Ken Howard, and Bob Hoffman.

Varying predicting algorithms depending on groups of flights

How much does route choice vary among airlines? Appendix D presented data dissected by airlines and the regions of the country. This was actually the first attempt in the analysis to look at the data (and, possibly, develop algorithms for route prediction) separately for different flight groups. We could all agree that airline and city-pair are strongly relevant to route prediction but how do the other characteristics rate? Are there any characteristics that we are not currently including that we should include, for instance, time of day or day of week?

Statistical methods to infer the relevant flight characteristics, such as regression, cluster analysis, factor analysis, and principle component analysis might help. (However, there is some skepticism towards factor analysis, because it might be as line fitting: with higher degree of a polynomial one may fit any set of points to it.) So, the task is to perform statistical analysis to determine the set of relevant flight characteristics and their relative values to route prediction.

As it was mentioned in Appendix D, determining predictability for groups of flights may help to measure confidence of Monitor/Alert data.

Another idea of flight grouping is to combine some aircraft types, i.e., by category (jets, turbo-props, etc.), by manufacture (Boeing, Airbus, etc.) or by aircraft weight. And yet another idea is to lump several city pairs together (i.e., SFO-EWR and SFO-JFK) in order to use substantial portions of their routes that are common. Suppose, we may determine a great portion on the SFO-JFK route by looking at the SFO-EWR route, then take an approach-to-JFK portion of some commonly used SFO-JFK routes, and then stitch them together. This idea reflects to an extent how traffic managers are thinking, how the Reroute Playbook is composed.

These ideas might not bring large improvements but improvements even in small percent of traffic may be critical if they fall on a specific traffic manager.

Definitely all these ideas belong to a completely new area and nobody knows how difficult it might be to deal with it. However, there is anecdotal evidence that some traffic managers already use uncertainty factor when evaluating the Monitor/Alert data: looking at some data they know from their experience that not all the flights are going to show up, and they mentally adjust the data.

Understanding limitations of the model

A possible way to improve predictability of default routes might lie in understanding the limitations of the model, rather than just its aggregate performance. Are there specific airlines, regions of the country, aircraft types, seasons, or days (e.g. convective weather) for which the model is particularly weak? Answering these questions will give some insight into a possible way to improve it and will give guidance to those using the model when they can trust it.

Inherent Predictability

How stable is the route choice between city-pairs over a space of a few hours: if it is not stable, then there doesn't seem to be much hope of improvement. Or may be it is not stable due to differences in traffic patterns as the day progresses? What impact does weather have on predictability?

Is there a way we can determine the inherent predictability of the data? That is, is there enough consistency in flight plans for same airline, same city-pair, same time period that we can even hope to have a good prediction? What is the reasonable limit of accuracy that we can even hope to achieve? Is the current method far off from this or close to it? This is a fundamental question that we maybe should have looked at first rather than just assuming that it can be much better, as everyone thinks it should be.

Getting internal airline knowledge

Getting more knowledge of what airlines are doing when selecting a route and incorporating this into analysis and ETMS might be extremely useful. It is better to avoid any route prediction model that tries to second guess what an airline is thinking when it chooses a flight plan, unless one has extensive conversations with the airlines first. This is too complicated and too heavily entangled in airline daily objectives that will always be unknown to us. Furthermore, one cannot even assume that the airlines will act in an optimal fashion – some airlines have. We are told that their flight planning software is much the same, and might even come from the same vendor, but they do not necessarily enter the same desired parameters or relative weights in multiple objective situations.

We have to talk to the airlines about how they select routes and mimic their methods as best we can. However, the problem with this approach is that airline dispatchers responsible for flight planning often talk on absolutely different language. Another problem is that even if ETMS tries to simulate behavior of major airlines only it might be too heavy a burden, and there are plenty of small carriers as well.

If we get inside source of what airlines are doing when selecting a route this will take us a lot farther than any statistical method. The objective might be to compute some weighted combination of the relevant information that they examine when making route choices. This would be a historical/statistical method.